



(12) **DEMANDE DE BREVET EUROPEEN**

(21) Numéro de dépôt: **93203620.5**

(51) Int. Cl.<sup>5</sup>: **H04N 7/18**

(22) Date de dépôt: **22.12.93**

(30) Priorité: **29.12.92 FR 9215836**

(43) Date de publication de la demande:  
**06.07.94 Bulletin 94/27**

(84) Etats contractants désignés:  
**DE FR GB IE**

(71) Demandeur: **LABORATOIRES  
 D'ELECTRONIQUE PHILIPS  
 22, Avenue Descartes  
 F-94450 Limeil-Brévannes(FR)**

(84) **FR**

(71) Demandeur: **PHILIPS ELECTRONICS N.V.  
 Groenewoudseweg 1  
 NL-5621 BA Eindhoven(NL)**

(84) **DE GB IE**

(72) Inventeur: **Lelong Pierre  
 Société Civile S.P.I.D.,  
 156 Boulevard Haussmann  
 F-75008 Paris(FR)**  
 Inventeur: **Dalm, Govert  
 Société Civile S.P.I.D.,  
 156 Boulevard Haussmann  
 F-75008 Paris(FR)**  
 Inventeur: **Klijn, Jan  
 Société Civile S.P.I.D.,  
 156 Boulevard Haussmann  
 F-75008 Paris(FR)**

(74) Mandataire: **Lottin, Claudine et al  
 Société Civile S.P.I.D.  
 156, Boulevard Haussmann  
 F-75008 Paris (FR)**

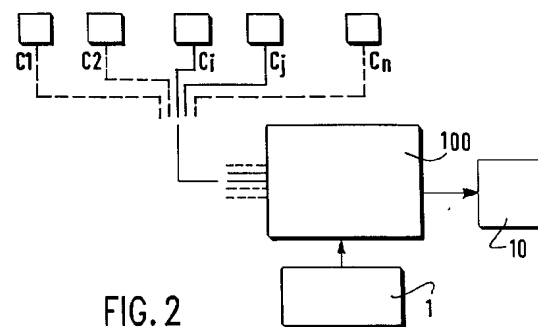
(54) **Procédé et dispositif de traitement d'images pour construire une image à partir d'images adjacentes.**

(57) L'invention concerne un procédé de traitement d'images, pour construire une image dite cible (I<sub>0</sub>), à partir d'images adjacentes, de cadrage fixe, dites images sources (I<sub>1</sub>,...,I<sub>j</sub>,...,I<sub>n</sub>), ayant des points de fuite sensiblement confondus. Ce procédé comprend des étapes de : numérisation des images ; détermination, pour un des pixels de l'image cible (I<sub>0</sub>), de l'adresse (A<sub>q</sub>) d'un point dit correspondant, dans une image source (I<sub>j</sub>) parmi toutes les images sources ; détermination de la valeur (F) de la luminance en ce point correspondant ; attribution au pixel de départ dans l'image cible (I<sub>0</sub>) de la valeur de luminance (F) de ce pixel correspondant ; et répétition de ces étapes pour chacun des pixels de l'image cible (I<sub>0</sub>).

L'invention concerne aussi un dispositif mettant en oeuvre ce procédé et comprenant un système de n caméras réelles fixes (C<sub>1</sub>,...,C<sub>n</sub>) qui fournissent n images sources (I<sub>1</sub>,...,I<sub>n</sub>) adjacentes couvrant un champ de vue de grand angle, et qui ont des points de fuite (P) communs ; et un système de construction d'image (100) simulant une caméra dite virtuelle (C<sub>0</sub>), mobile, pour fournir une sous-image dite cible

(I<sub>0</sub>), du champ de vue de grand angle, et construite à partir des images sources de même point de fuite (P).

Application : Dispositifs de surveillance.



L'invention concerne un procédé de traitement d'images, pour construire une image dite cible, à partir d'images adjacentes, de cadrage fixe, dites images sources, lesdites images cible et source ayant des points de fuite sensiblement confondus.

L'invention concerne également un dispositif de traitement d'images comprenant :

- un système de n caméras réelles fixes, disposées de manière à ce que leurs champs de vue individuels fusionnent pour former un seul champ de vue de grand angle pour l'observation d'une scène panoramique,
- un système de construction d'image simulant une caméra dite virtuelle, mobile, balayant continuellement la scène panoramique pour en fournir une sous-image dite cible, correspondant à une section arbitraire du champ de vue de grand angle, et construite à partir des images sources adjacentes fournies par les n caméras réelles, cette caméra virtuelle ayant un point de fuite confondu avec, ou proche de celui des caméras réelles.

L'invention trouve son application dans le domaine de la télésurveillance, ou bien dans le domaine de la télévision, lorsque des prises de vue couvrant des grands champs sont nécessaires par exemple dans le cas d'enregistrement de scènes de manifestations sportives. L'invention trouve aussi son application dans le domaine de la construction automobile pour réaliser des rétroviseurs périphériques et panoramiques sans angle mort.

Un dispositif de traitement d'image est déjà connu de la demande de brevet WO92-14341. Ce document décrit un système de traitement d'image appliqué à la télévision. Ce dispositif comprend une station émettrice incluant une pluralité de caméras fixes arrangées de manière adjacente pour que leurs champs de vue fusionnent et forment un champ de vue de grand angle. Ce système comprend aussi une station de traitement incluant des moyens pour générer un signal vidéo composite de la totalité de l'image correspondant au champ de vue de grand angle, et des moyens pour sélectionner une sous-image de cette image composite. Ce système comprend enfin des moyens, c'est-à-dire un moniteur, pour afficher cette sous-image. Cette sous-image correspond à un champ de vue d'angle inférieur à celui de l'image composite dit sous-section du champ de vue de grand angle.

Ce dispositif de traitement d'image est applicable uniquement aux systèmes de télévision classiques, où l'image est formée au moyen d'un faisceau de balayage ligne après ligne.

La station de traitement permet à un utilisateur d'opérer lui-même la sélection de la sous-section du champ de vue de grand angle. La sous-image correspondante a la même dimension que l'image fournie par une caméra individuelle. L'utilisateur

sélectionne cette sous-image en faisant varier le point de départ du balayage vis-à-vis de l'image composite correspondant au champ de vue de grand angle. Le champ de vue de grand angle a un axe parallèle au balayage vidéo, d'où il résulte que le point de départ du balayage vidéo de la sous-image peut être déplacé arbitrairement et continuellement parallèlement à cet axe.

Eventuellement, l'angle du champ de vue auquel la sous-image correspond peut être inférieur à celui d'une caméra réelle. Mais la localisation de la sous-image n'inclut pas de déplacement perpendiculairement au balayage ; sa localisation inclut seulement des déplacements parallèlement à ce balayage. Et la formation de cette sous-image n'inclut pas d'effet de zoom vis-à-vis de l'image composite, c'est-à-dire de changement de focale de la sous-image vis-à-vis de la focale des caméras de prise de vue.

Donc la station de traitement d'image comprend des moyens pour construire ligne après ligne la sous-image vidéo sélectionnée. Ces moyens incluent essentiellement un circuit de contrôle de la synchronisation des signaux vidéo des différentes caméras.

La présente invention a pour but de fournir un dispositif capable de simuler une caméra mobile, balayant le champ de vue de grand angle couvert par les n caméras fixes dont les champs de vue fusionnent.

Le but de la présente invention est tout particulièrement de fournir un tel dispositif simulant une caméra munie de toutes les facilités que pourraient avoir une caméra mobile existant réellement : c'est-à-dire à partir d'un observateur qui est fixe, des possibilités de déplacements angulaires horizontaux vers la gauche ou la droite d'une scène panoramique à observer ou à surveiller, des possibilités de déplacements angulaires verticaux vers le haut ou vers le bas de cette scène, des possibilités de rotation, et en outre des possibilités de zoom sur une partie de la surface de cette scène.

Ce but est atteint en mettant en oeuvre un procédé de traitement d'images, pour construire une image dite cible, à partir d'images adjacentes, de cadrage fixe, dites images sources, lesdites images cible et source ayant des points de fuite sensiblement confondus, caractérisé en ce qu'il comprend des étapes de :

- numérisation des images,
- détermination, pour un des pixels de l'image cible, de l'adresse d'un point dit correspondant, dans une image source parmi toutes les images sources,
- détermination de la valeur de la luminance en ce point correspondant,
- attribution au pixel de départ dans l'image cible de la valeur de luminance de ce pixel

correspondant,

- et répétition de ces étapes pour chacun des pixels de l'image cible.

Selon l'invention, pour mettre en oeuvre ce procédé, on propose en outre un dispositif de traitement d'images incluant :

- un système de n caméras réelles fixes, disposées de manière à ce que leurs champs de vue individuels fusionnent pour former un seul champ de vue de grand angle pour l'observation d'une scène panoramique,
- un système de construction d'image simulant une caméra dite virtuelle, mobile, balayant continuellement la scène panoramique pour en fournir une sous-image dite cible, correspondant à une section arbitraire du champ de vue de grand angle, et construite à partir des images dites sources adjacentes fournies par les n caméras réelles, cette caméra virtuelle ayant un point de fuite confondu avec, ou proche de celui des caméras réelles,

**caractérisé en ce que** ce système de traitement d'image est numérique et en ce que le système de construction (100) de l'image cible lo inclut :

- un calculateur d'adresse, pour faire correspondre à une adresse d'un pixel dans l'image cible, un point à une adresse dans une des images sources,
- des moyens pour calculer la valeur de luminance du point à l'adresse trouvée dans l'image source, et pour affecter cette valeur de luminance au pixel de départ à l'adresse dans l'image cible.

Ainsi, le dispositif selon l'invention permet de réaliser la construction d'une image cible comme celle que fournirait une caméra supposée, se déplaçant de manière continue ; cette image cible est formée à partir de plusieurs images sources adjacentes fournies chacune par une caméra du groupe de caméras disposées de manière fixe vis-à-vis de la scène à observer ; et, à partir de cette construction, ce dispositif peut fournir, par l'affichage sur écran, ou par enregistrement :

- soit une restitution séquentielle, image par image, de partitions de la scène observée, avec éventuellement effet de zoom,
- soit une restitution continue par balayage de la scène observée avec effet de site et d'azimut ou de rotation.

Dans une mise en oeuvre particulière, ce dispositif est caractérisé en ce que le système de reconstruction de l'image cible comprend :

- des premiers moyens de stockage des paramètres relatifs à la caméra virtuelle pour fournir au calculateur d'adresse le facteur d'échelle et l'orientation de l'axe optique de la caméra virtuelle dans un repère terrestre orthonormé fixe et indépendant des caméras,

définie par ses angles d'azimut, de site et de rotation ;

- des seconds moyens de stockage des paramètres relatifs aux caméras réelles pour fournir au calculateur d'adresse le facteur d'échelle et l'orientation de l'axe optique de chacune des caméras réelles définie par leurs angles d'azimut, de site et de rotation dans ledit repère terrestre fixe ;
- un générateur d'adresse pour générer les adresses (Ao) des pixels de l'image cible, pixel par pixel, de manière à couvrir la totalité de l'image cible, le calculateur d'adresse déterminant alors l'image source particulière et le point à l'adresse (Aq) dans cette image source, qui correspond à chaque pixel de l'image cible à partir des paramètres des caméras virtuelle et réelles.

Un autre problème technique se pose à la construction de l'image cible. En effet, on suppose que l'on dispose d'une pluralité de caméras positionnées de manière adjacentes et de telle sorte que le champ couvert par chacune ne laisse échapper aucune zone de la scène panoramique à construire : on suppose donc que l'on dispose de toutes les données pour construire l'image cible. Néanmoins, à chaque frontière entre les caméras, au passage d'une image d'une caméra, à une autre image d'une caméra adjacente, la différence d'angle de visée de ces deux caméras, pour deux zones adjacentes de la scène enregistrées par ces deux caméras différentes, crée de grandes distorsions d'image. Il en résulte que les partitions qui sont réalisées à cheval sur deux zones de la scène enregistrée par deux caméras différentes sont très inconfortables à visualiser et manquent totalement de précision.

Un autre but de l'invention est donc de fournir une telle construction d'image cible qui est corrigée en distorsion d'image à la frontière entre deux caméras, de telle sorte que cette ou ces frontières soit(ent) totalement invisible(s) par l'utilisateur.

Ce but est atteint au moyen d'un dispositif de traitement d'image tel que défini précédemment et en outre caractérisé ce que le calculateur d'adresse comporte :

- des premiers moyens de construction d'un modèle (MCo) de caméra virtuelle avec projection à travers le point de fuite,
- des seconds moyens de construction de modèles (MC1-MCn) de caméras réelles, avec projection à travers le point de fuite, et avec corrections des distorsions et défauts de perspective.

Dans une mise en oeuvre particulière, ce dispositif est caractérisé en ce que le calculateur d'adresse comporte :

- des premiers moyens de calcul de transformation géométrique pour appliquer à chaque pixel à une adresse (Ao) de l'image de la caméra virtuelle, une transformation géométrique dite "transformation perspective" inverse ( $H_0^{-1}$ ), qui prend en compte le modèle (MCo) de la caméra virtuelle fourni par les premiers moyens de construction, et les paramètres angulaires d'azimut, de site, de rotation et le facteur d'échelle de cette caméra virtuelle, fournis par les premiers moyens de stockage, pour déterminer à partir de cette transformation perspective inverse ( $H_0^{-1}$ ), le positionnement dans ledit repère terrestre, du rayon lumineux passant par ce pixel et le point de fuite,
- des moyens pour stocker le positionnement du rayon lumineux obtenu par la transformation perspective inverse ( $H_0^{-1}$ ),
- des moyens de sélection pour sélectionner l'image source particulière traversée par ce rayon lumineux,
- des seconds moyens de calcul de transformation géométrique pour appliquer à ce rayon lumineux dans ledit repère terrestre, une transformation géométrique dite "transformation perspective directe" ( $H_1-H_n$ ), qui prend en compte les modèles des caméras réelles fournis par les seconds moyens de construction, les paramètres angulaires d'azimut, de site, de rotation, et le facteur d'échelle de la caméra réelle correspondante, fournis par les seconds moyens de stockage,
- et des moyens de stockage pour fournir, à partir de cette transformation perspective directe ( $H_1-H_n$ ), l'adresse (Aq) dans l'image source particulière qui correspond au rayon lumineux, et donc au pixel d'adresse (Ao) dans l'image cible.

Avec ce dispositif, l'utilisateur qui surveille une scène panoramique obtient exactement le même confort, et les mêmes services qu'un utilisateur qui disposerait d'une caméra mobile, avec zoom, et avec des moyens mécaniques pour réaliser la variation de l'orientation de l'axe optique, c'est-à-dire pour réaliser des variations de site et d'azimut de la visée, ainsi que des rotations autour de l'axe optique de la caméra. L'avantage est que les moyens mécaniques sont évités. Ces moyens mécaniques, qui incluent des moteurs de rotation mécanique d'angles de site et d'azimut, et un moteur de réglage du zoom, sont toujours liés à des inconvénients : d'abord ils peuvent se bloquer, ensuite les déplacements engendrés sont très lents. De plus, il sont très onéreux. Enfin, étant le plus souvent installés en extérieur, ils se dégradent rapidement du fait des intempéries. Les moyens électroniques de traitement d'image utilisés selon

l'invention évitent tous ces inconvénients, car ils sont très précis, fiables, très rapides et faciles à commander. En outre, ils peuvent être disposés en intérieur, et ainsi à l'abri des intempéries. De plus, les moyens électroniques sont facilement programmables pour un fonctionnement automatique. Enfin, ils sont moins onéreux que les moyens mécaniques.

Ainsi l'utilisateur obtient, avec les moyens de l'invention, d'abord une image dépourvue de distorsions et ensuite des plus grandes précisions, et facilités d'exécution des manoeuvres de visée, que s'il disposait de moyens mécaniques. En outre, une scène panoramique de plus grand champ peut être observée, puisque l'on peut observer des champs de  $180^\circ$  ou même  $360^\circ$  selon le nombre de caméras utilisées. Par ailleurs, les manoeuvres peuvent être aisément programmées.

En matière de surveillance, il en résulte un grand progrès. Dans le domaine de la réalisation de rétroviseurs panoramiques pour automobiles, l'importance du progrès est aussi très grande.

Le fait d'utiliser plusieurs caméras pour faire la saisie des données nécessaires à la construction de l'image cible n'est pas un désavantage, car un tel ensemble de caméras fixes, à CCD (dispositif à couplage de charge), est devenu moins onéreux que ne le seraient les dispositifs mécaniques de variation de site, d'azimut et de rotation, ainsi que de zoom, associés à une seule caméra réelle mobile.

Dans une mise en oeuvre particulière ce système est caractérisé en ce que les moyens de détermination de luminance comprennent :

- un interpolateur pour calculer une valeur la plus probable d'une fonction de luminance (F) au point l'adresse (Aq) trouvée par le calculateur d'adresse dans l'image source fournie les moyens de sélection :
- des troisièmes moyens de stockage pour affecter la valeur (F) de luminance correspondant au point à l'adresse (Aq) trouvée dans l'image source, au pixel de départ dans l'image cible à l'adresse (Ao) fournie par le générateur d'adresse et en ce que le système de reconstruction de l'image cible comprend aussi :
- une interface pour permettre à un utilisateur de définir les paramètres de la caméra virtuelle, ces paramètres incluant son facteur d'échelle et l'orientation de son axe optique.

L'invention est décrite ci-après en détail, en référence avec les figures schématiques annexées dont :

- la FIG.1A représente une vue du dessus montrant les traces des différents plans images, dans le plan horizontal du repère terrestre, dans le cas où les caméras réelles ont

- des plans images perpendiculaires à ce plan horizontal ;
- la FIG. 1B représente le repère terrestre Px, Py, Pz, vu en projection dans le plan horizontal ;
  - la FIG.1C représente en vue plane, un plan image source avec son système d'axes de coordonnées particulier ;
  - la FIG.1D représente en vue plane, le plan image cible avec son système d'axes de coordonnées particulier ;
  - la FIG. 1E représente l'effet de la délimitation d'une section du champ de vue de grand angle de deux caméras réelles adjacentes, au moyen des paramètres choisis par l'utilisateur pour la caméra virtuelle, pour construire une sous-image d'une scène panoramique ;
  - la FIG.1F représente l'image cible construite par la caméra virtuelle définie par ces paramètres, cette image cible étant composée d'une première partie d'image construite à partir de l'image source fournie par la première des deux caméras réelles, et d'une seconde partie d'image construite à partir de l'image source fournie par la seconde de ces caméras ;
  - la FIG.1G représente une disposition relative de trois caméras réelles adjacentes pour couvrir un champ de vue de 180° ;
  - la FIG.2 qui représente en blocs fonctionnels le dispositif de traitement d'image avec le système de construction de l'image cible, les caméras réelles, l'interface-utilisateur et le système de visualisation de l'image cible ;
  - la FIG.3 qui représente le dispositif de traitement d'image sous forme de blocs fonctionnels plus détaillés que la FIG.2 ;
  - la FIG.4 qui illustre le calcul d'une valeur d'une fonction de luminance relativement à une adresse dans une image source ;
  - la FIG. 5A qui illustre la modélisation des caméras réelle et virtuelle ;
  - la FIG. 5B qui illustre en projection sur le plan horizontal du repère terrestre, l'effet des perspective et distorsion sur les positions des points correspondants de même luminance dans l'image cible et dans l'image source traversées par un même rayon lumineux passant par ces points ;
  - la FIG.6 qui représente, sous forme de blocs fonctionnels, le calculateur d'adresse qui calcule l'adresse du point dans l'image source, correspond à un pixel à une adresse donnée dans l'image cible ;
  - la FIG.7A qui représente une première image source numérique telle que formée par une première caméra fixe réelle et la FIG.7B qui représente une seconde image source telle

que formée par une seconde caméra fixe réelle adjacente à la première

- la FIG.7C représente une image cible numérique reconstruite de la même façon que dans le cas de la FIG. 1F, sur laquelle est mise en évidence les défauts de distorsion et perspective entre la première partie d'image cible construite à partir de la première image source, et la seconde partie d'image cible construite à partir de la seconde image source ;
- la FIG.7D représente l'image cible numérique de la FIG.7C après traitement par le dispositif de traitement d'image où les défauts de distorsion et perspective sont éliminés.

#### I/ Le système de prise de vue

La Fig.1G montre une disposition possible de plusieurs caméras fixes réelles pour enregistrer les données relatives à une scène de 180° d'angle. L'enregistrement de cette scène panoramique est fait avec trois caméras fixes C1, C2, C3. Les champs optiques des caméras sont tels qu'absolument tous les détails de la scène panoramique sont enregistrés par une caméra ou par une autre, de manière à ne laisser échapper aucun objet de la surveillance. Les caméras sont disposées de manière à avoir un point de fuite commun P (en anglais : VIEW POINT) ou des points de fuite très proches.

Les axes PZ1, PZ2, PZ3 représentent respectivement les axes optiques des caméras C1, C2, C3 ; et les points O1, O2, O3 représentent respectivement les centres géométriques des images I1, I2, I3, dans les plans images, sur les axes optiques.

On pourrait effectuer une surveillance horizontale sur 360° en disposant de manière appropriée 6 caméras fixes. Mais on pourrait aussi effectuer une surveillance verticale, ou dans les deux directions. L'homme du métier peut, sans description plus approfondie des diverses dispositions des caméras fixes les unes vis-à-vis des autres, réaliser toute sorte d'observation d'une scène panoramique.

En référence avec la FIG. 1A, le dispositif de prise de vue comprend une pluralité de n caméras réelles fixes, de focales fixes et connues, disposées de manière adjacente, et de telle sorte que leurs champs de vue individuels fusionnent pour couvrir un champ de vue de grand angle. Ainsi, les n caméras fixes adjacentes fournissent n images fixes adjacentes, de sorte que ce dispositif de prise de vue peut surveiller une scène panoramique. Les champs optiques de caméras sont tels que tous les détails de la scène panoramique sont enregistrés par une caméra ou par une autre, pour ne laisser échapper aucun objet de la surveillance.

Pour obtenir ce résultat, ces  $n$  caméras fixes adjacentes sont en outre disposées de telle sorte que leurs centres optiques  $P$ , encore appelés points de fuite (ou VIEW POINT en anglais) coïncident. Le point de fuite d'une caméra est défini comme le point tel que tout rayon issu d'une source lumineuse et passant par ce point, traverse l'optique de la caméra sans déviation.

En fait, les points de fuite des  $n$  caméras ne peuvent pas physiquement coïncider. Mais on considère ci-après que la condition de coïncidence est suffisamment bien remplie si la distance qui sépare chacun de ces points de fuite les uns des autres est petite vis-à-vis de leur distance à la scène panoramique filmée, par exemple si leur distance respective est 5 cm ou 10 cm, et la distance à la scène panoramique est 5 m. Ainsi la condition de coïncidence est estimée remplie si le rapport de ces distances est de l'ordre de, ou supérieur à, 50, et il n'est pas besoin, selon l'invention, d'utiliser des optiques à miroirs difficiles à ajuster, et coûteuses, pour arriver à une coïncidence rigoureuse des points de fuite.

## II/ La formation des images par les caméras

Le but de l'invention est de fournir un système de reconstruction d'image numérique qui simule une caméra mobile capable, avec des réglages choisis par un utilisateur, de fournir une image numérique de n'importe quelle partie, ou sous-image, de la scène panoramique enregistrée par les  $n$  caméras fixes.

Les  $n$  caméras sont numérotées  $C1.... Ci, Cj... Cn$  ; ces caméras fournissent respectivement des images dites sources numériques  $I1..., li, lj..., In$ . Dans ce qui suit, on considérera par exemple les images sources  $li$  et  $lj$  formées par deux caméras réelles fixes adjacentes  $Ci$  et  $Cj$ .

Ces caméras  $Ci$  et  $Cj$  réelles fixes forment des images respectives de la scène panoramique dans des plans images sources adjacents  $li$  et  $lj$ . Sur la FIG. 1A, les axes  $PZi$  et  $PZj$ , passant respectivement par les centres géométriques  $Oi$  et  $Oj$  des images sources  $li$  et  $lj$ , représentent les axes optiques des caméras réelles fixes  $Ci$  et  $Cj$ .

En référence avec la FIG. 1B, on définit d'abord un repère terrestre  $Px, Py, Pz$ , d'axes orthogonaux dans lequel les axes  $Px$  et  $Pz$  sont horizontaux, et l'axe  $Py$  est vertical.

Les images sources, telles que les images  $li$  et  $lj$ , sont numérisées, et chaque pixel  $m$  de ces images est repéré par ses coordonnées dans le plan image. A cet effet, comme représenté sur la FIG.1C, on définit dans chaque plan image un repère de coordonnées rectangulaires  $(OiXi, OiYi)$  et  $(OjXj, OjYj)$  dans lequel les axes  $OiXi$ , ou  $OjXj$ , sont horizontaux, c'est-à-dire dans le plan du repè-

re terrestre  $Px, Pz$ . Les plans images définis par  $(OiXi, OiYi)$  et  $(OjXj, OjYj)$  sont perpendiculaires aux axes optiques  $PZi$ , et  $PZj$  et ont pour centres géométriques  $Oi$  et  $Oj$  respectivement.

Une fois établis les repères individuels liés à chaque plan image des caméras, ces plans images sources fixes peuvent être liés au repère terrestre par :

- leur angle d'azimut (en anglais PAN)  $\theta_i, \theta_j$ ,
- leur angle de site (en anglais TILT)  $\Phi_i, \Phi_j$ .

L'angle d'azimut  $\theta_i$  ou  $\theta_j$ , est l'angle que fait un plan vertical contenant l'axe optique  $PZi$  ou  $PZj$ , avec l'axe horizontal  $Pz$  du repère terrestre. C'est donc un angle de rotation horizontale autour de l'axe vertical  $Py$ .

L'angle de site  $\theta_i$  ou  $\theta_j$  est l'angle que fait l'axe optique  $PZi$  ou  $PZj$  avec le plan horizontal  $(Px, Pz)$ . C'est donc un angle de rotation verticale autour d'un axe horizontal, l'axe  $OiXi$  ou  $OjXj$  de chaque plan image.

Dans l'exposé ci-après, pour des raisons de simplicité, on a supposé, à titre d'exemple, en référence avec la FIG.1A que les plans images  $li, lj$  sources fournis par les caméras fixes  $Ci, Cj$  étaient verticaux, c'est-à-dire que leurs angles de site  $\Phi_i, \Phi_j$  étaient nuls.

Pour des raisons de simplicité également, on a noté par la même référence sur la FIG. 1A la trace des plans et des axes, et les plans et axes correspondants, aussi bien pour les images sources que pour l'image cible décrite plus loin.

La FIG.1A, qui est une vue du dessus schématique des images formées, laisse donc apparaître seulement les traces  $li$  et  $lj$  des plans images sources fixes représentées par des segments dans le plan horizontal  $Px, Pz$ .

La FIG.1E montre par exemple les images  $li$  et  $lj$  contigües de la scène panoramique fournies par deux caméras fixes  $Ci$  et  $Cj$  adjacentes. Sur la FIG. 1E, pour la simplicité de l'exposé, les images  $li$  et  $lj$  sont représentées toutes les deux projetées dans un même plan, alors que dans la réalité, ces images font entre elles un angle égal à celui des axes optiques des caméras fixes. Dans ces images, l'utilisateur peut choisir de regarder n'importe quelle sous-image délimitée par la ligne  $Jo$ , plus ou moins à gauche ou à droite, plus ou moins en haut ou en bas, avec le même grandissement que les caméras fixes, ou avec un grandissement supérieur, ou éventuellement avec un grandissement inférieur.

La caméra mobile simulée est capable de construire une image dite cible  $lo$  à partir des parties d'image source  $Si, Sj$  délimitées par la ligne  $Jo$  de la FIG. 1E. Cette caméra référencée  $Co$  ci-après est appelée caméra virtuelle du fait qu'elle simule une caméra non réellement existante. Evidemment cette caméra mobile n'est pas limitée à

explorer les deux images  $li$ ,  $lj$ . Elle peut explorer toutes les images sources de  $l1$  à  $ln$ .

Cette caméra virtuelle  $Co$  peut être définie de la même façon que les caméras réelles fixes par :

- son angle d'azimut  $\theta_o$
- son angle de site  $\phi_o$
- son angle de rotation  $\psi_o$
- et son grandissement (effet de zoom) concrétisé par sa longueur focale  $PO_o$ , notée  $z_o$ ,

à condition de choisir son point de fuite  $P$  confondu avec les points de fuite  $P$  des caméras réelles fixes,  $O_o$  étant le centre géométrique de l'image cible  $lo$ . Le point de fuite de la caméra virtuelle est confondu avec le point de fuite approximatif défini plus haut pour les caméras réelles.

La FIG. 1A montre la trace notée  $lo$  du plan image de la caméra virtuelle dans le plan horizontal et son axe optique  $PZ_o$  passant par le centre géométrique  $O_o$  de l'image cible  $lo$ .

Dans la définition de cette caméra virtuelle  $Co$  mobile, l'angle d'azimut  $\theta_o$  est l'angle que fait le plan vertical contenant son axe optique  $PZ_o$  avec l'axe  $Pz$  horizontal du repère terrestre ; l'angle de site  $\phi_o$  est l'angle que fait son axe optique  $PZ_o$  avec le plan horizontal  $Px$ ,  $Pz$  du repère terrestre ; son angle  $\psi_o$  est l'angle de rotation de la caméra virtuelle autour de son propre axe optique, ce dernier étant fixe ; enfin, sa focale  $PO_o$  est variable, ce qui permet de changer le grandissement de cette image cible par rapport à celui des images sources (effet de ZOOM).

Ainsi, en faisant varier les angles d'azimut et de site,  $\theta_o$  et  $\phi_o$ , son angle de rotation  $\psi_o$  et la focale  $PO_o$ , la caméra virtuelle est tout à fait assimilable à une caméra mobile qui balaye le champ de vue de grand angle formé par les champs de vue fusionnés des différentes caméras réelles fixes  $C1$  à  $Cn$ .

On remarquera que la caméra virtuelle  $Co$  peut viser une partie petite (ou sous-section), délimitée par  $Jo$ , du champ de vue de grand angle et en réaliser une image agrandie  $lo$  par exemple de même dimension finale que chacune des images  $l1...ln$  fournie par chacune des caméras réelles  $C1...Cn$ , en jouant sur sa focale variable  $PO_o$ .

On remarquera également que le déplacement du champ de vue de la caméra mobile  $Co$  peut être continu et arbitraire ; ce champ de vue correspondant à  $Jo$  peut se trouver à cheval sur deux parties ( $Si, Sj$ ) d'image  $li$  et  $lj$  contigües en  $LO$ , fournies par deux caméras adjacentes  $Ci$  et  $Cj$ .

Dans ce cas, l'image  $lo$  construite par la caméra virtuelle  $Co$  contient deux parties d'image différentes, l'une  $loi$  étant construite à partir des informations  $Si$  contenues dans l'image numérique  $li$ , et l'autre  $loj$  étant construite à partir des informations  $Sj$  contenues dans l'image numérique  $lj$ . Dans la FIG.1A,  $loi$  et  $loj$  représentent les traces des ima-

ges cibles  $loi$  et  $loj$  dans le plan horizontal.

Comme pour les images réelles, on définit maintenant, en référence avec la FIG.1D, dans le plan de l'image cible numérique  $lo$ , un repère de coordonnées rectangulaires ( $O_o X_o$ ,  $O_o Y_o$ ) dans lequel l'axe  $O_o X_o$  est horizontal, c'est-à-dire dans le plan horizontal du repère terrestre  $Px$ ,  $Pz$ . Le pixel  $O_o$  étant le centre géométrique de l'image cible  $lo$ , en outre situé sur l'axe optique  $PZ_o$  de la caméra virtuelle  $Co$ . Chaque pixel  $m'$  du plan image cible  $lo$  est donc repéré par ses coordonnées dans ce système d'axes rectangulaires ( $O_o X_o$ ,  $O_o Y_o$ ).

On définit également un cadrage de l'image cible, c'est-à-dire ses dimensions en nombre de pixels selon la direction horizontale et nombre de lignes selon la direction verticale.

### III/ La méthode de construction de l'image cible à partir des images sources

La tâche des moyens de traitement d'images numériques selon l'invention est donc la construction de "l'image cible" fournie par la caméra virtuelle, à partir des "images sources" fournies par les caméras réelles fixes.

On a vu que, dans la construction de l'image cible  $lo$ , se posait un problème technique parce que les images  $li$  et  $lj$  fournies par les caméras réelles adjacentes, telles que représentées sous forme d'images numériques de paysage sur les FIG.7A et 7B, respectivement montrent des distorsions qui font qu'elles ne se raccordent pas correctement : en particulier certaines parties de lignes droites sont courbes, les verticales ne sont pas verticales etc. Ce qui fait que, au raccord, ces lignes vont être sécantes au lieu d'être dans le prolongement les unes des autres. De plus, les parties d'images sources de part et d'autre du raccord proviennent de parties de la scène panoramique à observer qui sont vues sous des perspectives différentes. D'où il résulte que les parties d'image cible  $loi$  et  $loj$  pourraient être mal raccordées, comme il apparaît dans l'image numérique cible représentée à titre d'exemple sur la FIG.7C et formée par la juxtaposition pure et simple des parties des images numériques des FIG.7A et 7B.

La présente invention propose une méthode et des moyens pour s'affranchir de ces défauts et fournir une image numérique reconstruite qui est dépourvue des défauts de distorsions et de perspectives, et dont les parties  $loi$ ,  $loj$  qui la composent se raccordent de manière telle que la droite frontière est invisible pour l'observateur et utilisateur.

La FIG.7D montre l'image numérique du paysage de la FIG.7C dans laquelle les défauts de distorsion et perspective ont été corrigés par les

moyens de l'invention.

La méthode générale de construction de l'image source comprend différentes étapes qui sont mises en oeuvre par le dispositif de traitement de signal décrit plus loin.

Cette méthode comprend d'abord une étape où :

- à chaque pixel  $m'$  ayant une adresse  $A_o$  dans l'image cible  $l_o$ , on fait correspondre un point  $m$  à une adresse  $A_q$  dans une image source, cette adresse étant définie par :
- la référence ou index de la caméra réelle qui va fournir l'image source ;
- l'adresse  $A_q$  du point  $m$  dans cette image source. Cette méthode comprend alors une seconde étape où :
- la valeur la plus probable de la luminance est évaluée en ce point  $m$  de l'image source,
- puis cette valeur de luminance est affectée au pixel  $m'$  dans l'image cible.

Ces étapes sont menées pour tous les pixels  $m'$  de l'image cible  $l_o$ .

Cependant, les moyens de traitement doivent rendre à l'image cible construite, toutes les qualités d'une image que l'observateur obtiendrait avec une caméra mobile conventionnelle :

- absence de distorsions, ajustage des perspectives,
- absence de droite de coupure à la frontière entre deux ou plusieurs images adjacentes.

Le problème est donc de rendre ces droites frontières invisibles.

#### IV/ Le dispositif de traitement d'image dans ses éléments essentiels

La Fig.2 montre sous forme de blocs fonctionnels les différents éléments du dispositif de traitement d'image selon l'invention.

Les blocs  $C_1, C_i, C_j, \dots, C_n$  représentent les  $n$  caméras fixes réelles dont les sorties fournissent les  $n$  images sources notées  $l_1, \dots, l_i, l_j, \dots, l_n$ .

Chacune des caméras réelles a des paramètres de prise de vue correspondant à son orientation par rapport à la scène, ainsi qu'un facteur d'échelle. Ces paramètres doivent être connus de manière très précise.

Le bloc 1 représente un système de contrôle qui permet à l'utilisateur de choisir et d'afficher des paramètres relatifs à l'orientation et à un facteur d'échelle de la caméra virtuelle  $C_o$ .

Le bloc 100 est un système de reconstruction d'image qui permet de calculer, à partir des images sources  $l_1, \dots, l_n$  fournies par les caméras réelles  $C_1, \dots, C_n$ , l'image dite "image cible"  $l_o$  qui serait donnée par la caméra virtuelle  $C_o$  orientée et réglée avec les paramètres définis par le système de contrôle 1, cette caméra virtuelle  $C_o$  étant dispo-

sée de telle sorte que son point de fuite est confondu ou proche du point de fuite  $P$  des caméras réelles.

Le bloc 10 correspond à des moyens d'affichage, ou bien d'enregistrement, qui peuvent être entre autres, un moniteur d'affichage numérique pour l'affichage en temps réel et/ou un magnétoscope pour l'enregistrement sur bandes magnétiques et...

D'autre part, il faut noter que les caméras réelles peuvent fournir des données analogiques. Dans ce cas, des modules de conversion analogique-numérique non représentés sont utilisés pour former les images sources numériques.

D'une façon avantageuse, l'homme du métier peut choisir des caméras à couplage de charges (CCD, de l'anglais Charge Coupled Device). Ce genre de caméra est peu onéreux, léger, robuste et fiable. Sa résolution est très bonne, et peut encore augmenter dans le futur en fonction de l'évolution des technologies.

#### V/ Le dispositif de traitement d'image dans ses éléments détaillés

La Fig.3 montre un diagramme de blocs fonctionnels correspondant aux moyens essentiels pour réaliser le traitement.

Ces moyens comportent principalement :

- une interface utilisateur 2,
- un calculateur d'adresse 200, incluant un sélecteur d'image source 205 ;
- un interpolateur 112.
- un générateur d'adresse 113 qui génère les adresses  $A_o$  des pixels de l'image cible  $l_o$ , pixel par pixel, de manière à couvrir la totalité de cette image cible  $l_o$ ,
- des premiers moyens de stockage 210 des paramètres  $\phi_o, \theta_o, \psi_o$  et  $z_o$  (le facteur d'échelle) de la caméra virtuelle  $C_o$ ,
- des seconds moyens de stockage 21.1 à 21.n des paramètres  $\phi_1$  à  $\phi_n, \theta_1$  à  $\theta_n, \psi_1$  à  $\psi_n$  et  $z_1$  à  $z_n$  (les facteurs d'échelles des caméras réelles  $C_1$  à  $C_n$ ) respectivement,
- des troisièmes moyens de stockage 117.

Pour chaque pixel  $m'$  de l'image cible  $l_o$  à construire, une adresse  $A_o$  est stockée dans le bloc 113 générateur des adresses dans l'image cible  $l_o$ . En fonction des réglages choisis par l'utilisateur qui sont entrés dans le bloc de contrôle 1 au moyen d'une interface-utilisateur proprement dite 2, et qui sont stockés sous la forme des paramètres  $\phi_o, \theta_o, \psi_o$  et  $z_o$  de la caméra virtuelle dans le bloc 210, et en fonction des paramètres  $\phi_i, \theta_i, \psi_i$  et  $z_j$  ( $j$  représentant ici l'indice de toutes les caméras de 1 à  $n$ ) des caméras réelles qui sont stockés dans les blocs 21.1, 21.j, 21.n, le calculateur d'adresse 200 fournit l'index  $j$  (parmi tous les indices 1 à  $n$ ) de la caméra réelle  $C_j$  capable de



fournir l'image Ij où trouver les données correspondant à l'adresse Ao du pixel m' de l'image cible Io.

Alors le sélecteur d'image 205 sélectionne cette caméra réelle Cj parmi les n caméras réelles, et fournit l'image source Ij correspondante où trouver ces données à traiter.

Le calculateur d'adresse 200, à partir des mêmes éléments, calcule l'adresse Aq du point m correspondant dans cette image source Ij sélectionnée, et ceci à travers les différentes transformations ou perspectives imposées par l'utilisateur en 2.

Par ces opérateurs, on établit donc une correspondance entre un pixel m' à une adresse donnée Ao de l'image cible Io, et un point m à une adresse calculée Aq d'une image source Ii.

Puis, la valeur de la luminance au point m à l'adresse Aq dans l'image source Ij est déterminée, afin d'être affectée au pixel correspondant m' à l'adresse Ao dans l'image cible Io. Et ainsi pour tous les pixels de l'image cible Io.

Mais, alors que l'adresse Ao dans l'image cible Io est réellement celle d'un pixel m', à la différence, l'adresse Aq dans l'image source Ii est précisément une adresse calculée, ce qui fait que, en général, cette adresse trouvée ne coïncide pas avec un pixel, mais avec un point m disposé entre des pixels. La luminance de ce point m doit donc être également calculée.

La Fig.4 illustre le problème qui se pose lorsque l'adresse Aq ne coïncide avec aucun pixel d'une image source Ii.

Ce problème est traité par l'interpolateur 112.

L'interpolateur 112 prend en compte l'intensité lumineuse, c'est-à-dire par exemple les niveaux de gris, ou une valeur d'une fonction de luminance des pixels d'un voisinage de l'adresse Aq, pour calculer l'intensité lumineuse à affecter à l'adresse Aq en question. Et par une méthode d'interpolation, le bloc 112 calcule l'intensité la plus probable à affecter à l'adresse Aq.

Il existe plusieurs fonctions possibles d'interpolation qui sont plus ou moins complexes, et qui fournissent selon leur complexité, des résultats plus ou moins finement ajustés.

Une fonction d'interpolation très simple est l'interpolation bilinéaire illustrée par la Fig.4. L'adresse Aq comporte une partie entière et une partie fractionnaire. L'adresse Aq se trouve pour cette raison située entre quatre pixels voisins m1, m2, m3, m4. La partie fractionnaire de l'adresse est représentée par  $\Delta x$  compté positivement de m1 vers m2, et par  $\Delta y$  compté positivement de m1 vers m3. La distance entre m1 et m3, et entre m1 et m2 est de 1 pixel.

L'intensité F ou valeur de la fonction de luminance à l'adresse Aq est calculée à partir des valeurs de  $\Delta x$ ,  $\Delta y$  et des intensités (ou valeurs des

fonctions de luminance) des pixels m1, m2, m3, m4 du voisinage, soit respectivement F1, F2, F3, F4

$$F(Aq) = F1(1-\Delta x)(1-\Delta y) + F2\Delta x(1-\Delta y) + F3(1-\Delta x)\Delta y + F4\Delta x\Delta y$$

L'intensité F(Aq) calculée par l'interpolateur à l'adresse Aq dans les images sources, est ensuite attribuée comme intensité (ou valeur de fonction de luminance) au pixel m' situé à l'adresse de départ Ao dans l'image cible Io, et stockée dans le bloc 117.

## VI/ Description du calculateur d'adresse 200

En référence avec la FIG.6, le calculateur d'adresse 200 est décrit ci-après en détail.

Le calculateur d'adresse 200 prend d'abord en compte les paramètres des caméras virtuelles et réelles telle que définis précédemment.

Le paramètres  $\phi_0$ ,  $\theta_0$ ,  $\psi_0$  et  $z_0$  de la caméra virtuelle sont disponibles dans le module de stockage 210 ; les paramètres  $\phi_1$  à  $\phi_n$ ,  $\theta_1$  à  $\theta_n$ ,  $\psi_1$  à  $\psi_n$  et  $z_1$  à  $z_n$  des caméras réelles sont disponibles respectivement dans les modules de stockage 21.1 à 21.n.

On a déterminé une définition des dimensions des images en nombre de pixels qui s'exprime par largeur x hauteur (en pixels). Donc les facteurs  $z_0$  ou  $z_i$  d'échelle s'expriment en pixels/mm. Les angles sont exprimés par exemple en degrés.

Pour la caméra virtuelle, ces paramètres  $\phi_0$ ,  $\theta_0$ ,  $\psi_0$ ,  $z_0$  sont calculés à partir de l'orientation ou positionnement de l'axe optique de la caméra virtuelle, du positionnement de son point de fuite P et centre d'image Oo, et du facteur d'échelle choisis par l'utilisateur, entrés sur l'interface 2.

Pour les caméras réelles, les paramètres  $\phi_1$  à  $\phi_n$ ,  $\theta_1$  à  $\theta_n$ ,  $\psi_1$  à  $\psi_n$ , les distances focales PO1 à PO<sub>n</sub>, les facteurs d'échelle  $z_1$  à  $z_n$  et le centre O1 à On de l'image numérique sont déterminés de façon très précise, à partir de calibrages et mesures très précises effectuées sur le positionnement du point P et centre O1 à On de chaque caméra, son facteur d'échelle, et l'orientation ou positionnement de son axe optique.

Le calculateur d'adresse 200 comprend ensuite des moyens 201 et 206 pour construire une modélisation de la caméra virtuelle Co, et une modélisation des caméras réelles C1, C2, C3,...Cn.

La Fig.5A illustre le principe de modélisation des caméras virtuelle Co, et réelles C1, C2... Cn. Le modèle de projection représenté sur la Fig.5A est fondé sur une projection conique à travers un point selon laquelle tous les rayons optiques passent par le seul point de fuite P (en anglais PIN-HOLE MODEL). Le point de fuite P (en anglais :

VIEW POINT) est supposé confondu pour toutes les caméras, réelles ou virtuelle, sous les conditions exposées précédemment.

Le repère terrestre  $P_x, P_y, P_z$ , et le repère individuel de la caméra virtuelle sont définis dans le module de construction 201 ; le repère terrestre  $P_x, P_y, P_z$  et les repères individuels de chacune des caméras réelles  $C_1$  à  $C_n$  sont définis dans le module de construction 206.

Supposons, en référence avec la Fig.5A que  $M$  soit un point objet situé dans une scène à surveiller. Le point  $m'$  correspondant à  $M$  dans le plan image de la caméra virtuelle  $Co$  est situé à l'intersection de la droite  $PM$  avec le plan image  $lo$  (Ooxo, Ooyo) de la caméra  $Co$ .

Il faut remarquer que le modèle de la FIG.5A est directement valable pour la caméra virtuelle qui est une caméra idéale sans distorsions. Lorsque l'on a affaire à une caméra réelle, par exemple munie d'un objectif à grand angle d'ouverture, il existe dans l'image des distorsions qu'il faut corriger.

La FIG.5B illustre ce problème. Cette figure 5B montre les traces des plans image cible  $lo$ , et image source  $lj$  représentés comme dans la FIG.1A.

Le problème à résoudre, pour construire l'image cible  $lo$  à partir des images sources  $li$  à  $ln$  est de déterminer à quelle adresse  $Aq$  on peut trouver le point  $m$  qui correspond exactement au pixel  $m'$  d'adresse  $Ao$ . On cherche en effet à déterminer l'intensité qui existe dans l'image source en ce point  $m$ , pour pouvoir l'affecter au point  $m'$  dans l'image cible.

Cette détermination n'est pas simple du fait des défauts de distorsions et défauts de perspective qui affectent les caméras réelles.

En référence avec la FIG.5B, on a représenté comme dans la FIG.5A que le pixel  $m'$  d'adresse  $Ao$  se trouve à l'intersection du rayon lumineux  $PM$  avec l'image cible  $lo$ ,  $PM$  étant le rayon qui joint le point de fuite  $P$  à un point  $M$  de la scène à observer puisque, dans cette caméra idéale, tout rayon passant par le point de fuite  $P$  traverse le plan image sans déviation.

Par ailleurs, ce rayon  $PM$  coupe le plan image source  $lj$  en un point  $m^*$ . La caméra  $C_j$  sera donc la caméra réelle sélectionnée pour fournir les données nécessaires à la construction de cette partie de l'image cible.

Mais par contre, il n'est pas vrai que le point  $m^*$  situé à l'intersection entre le même rayon  $PM$  et le plan image source  $lj$  soit le point correspondant de  $m'$  dans  $lj$  capable de fournir la valeur d'intensité  $F$  à affecter à ce point  $m'$ .

Il faut prendre en compte que la caméra  $C_j$ , comme toute caméra, comprend un objectif optique, qui forme une image réelle sur un substrat

muni d'une matrice de CCD ; cette image réelle est ensuite transformée par un module de conversion analogique-numérique pour fournir l'image source numérique  $lj$  qui est alors traitée par le dispositif de traitement d'image numérique selon l'invention. Du fait de la nécessité d'utiliser un objectif à lentille, on peut considérer, comme il est montré schématiquement sur la FIG.5B, que le rayon lumineux  $MP$  est dévié par le champ de distorsion  $DF$  représenté en traits pointillés, ce qui fait que le point correspondant à  $m'$  n'est pas le point  $m^*$  situé sur  $MP$ , mais bien un point  $m$  situé à une petite distance de  $m^*$  dans le plan image source  $lj$ .

Avec les données relatives aux repères terrestre et individuel, le module de construction 201 bâtit un modèle  $MCo$  de caméra virtuelle sur le modèle de projection conique représenté sur la FIG.5A, dans lequel tout rayon  $PM$  dans l'espace à 3 dimensions correspond à un pixel  $m'$  dans l'espace à 2 dimensions constitué par le plan image cible  $lo$ , puisque la caméra virtuelle  $Co$  est considérée comme une caméra idéale.

D'autre part, avec les données relatives au repère terrestre et aux repères individuels, le module de construction 206 bâtit les modèles  $MC_1$  à  $MC_n$  relatifs à chacune des caméras réelles, sur le modèle de projection conique représenté sur la FIG.5A, dans lequel tout rayon  $PM$  dans l'espace à 3 dimensions correspond à un point  $m^*$  dans l'espace à 2 dimensions constitué par le plan image source de chacune des caméras. Et, comme ces caméras ne sont pas des caméras idéales, mais présentent au contraire des distorsions, ce module 206 de construction des modèles  $MC_1$  à  $MC_n$  contient des moyens de corrections de distorsions qui permettent de remplacer systématiquement le point  $m^*$  trouvé, qui correspond à une image source distordue, par un point  $m$  dans le même espace à deux dimensions, qui correspond à une image source corrigée. Les moyens de correction des distorsions des objectifs peuvent être tous moyens connus de l'homme du métier.

Le calculateur d'adresse 200 comprend en outre des moyens de calcul d'adresse proprement dits pour déterminer l'adresse  $Aq$  de ce point  $m$  de l'image source distordue  $lj$ .

Ces moyens sont représentés sous forme de blocs fonctionnels sur la FIG.6, et mettent en oeuvre une méthode du calcul de l'adresse  $Aq$ , décrite ci-après, et illustrée par la FIG.5B pour la clarté de l'exposé.

#### VI a) La méthode de calcul d'adresse

Cette méthode de calcul de l'adresse  $Aq$  correspondant à une adresse de départ  $Ao$ , comporte plusieurs étapes qui incluent au moins une transformation géométrique complexe qui est le produit

de plusieurs transformations élémentaires. Cette transformation géométrique est fondée sur le fait qu'il existe maintenant, grâce au module 206 décrit précédemment, une correspondance entre chaque pixel  $m$  d'une image source d'une caméra réelle et un rayon lumineux PM, dans l'espace à trois dimensions, passant par le point de fuite  $P$  défini dans le modèle de la caméra. La localisation d'un pixel  $m$  dans l'image dépend des paramètres de la caméra relatifs à la définition du modèle de cette caméra, alors que le rayon lumineux correspondant PM a une localisation dans l'espace à trois dimensions qui peut être exprimée par rapport au "repère terrestre" fixe.

La méthode de calcul de l'adresse  $A_q$  à partir de  $A_o$  inclut les étapes suivantes :

- une première transformation dite "transformation-perspective inverse" notée  $H_o^{-1}$ , qui donne la direction du rayon lumineux PM correspondant à l'adresse de départ  $A_o$  dans l'image cible  $I_o$ ,
- la recherche de l'index  $j$  la caméra réelle  $C_j$  dont le champ de visée contient le rayon lumineux en question, et si plusieurs caméras sont concernées, du fait d'un recouvrement possible des champs de visée, le choix d'une caméra appropriée,
- une seconde transformation dite "transformation-perspective directe" notée  $H_j$ , qui donne l'adresse  $A_q$  dans l'image  $I_j$  de la caméra réelle  $C_j$  sélectionnée, correspondant au rayon lumineux PM.

Les deux transformations, inverse  $H^{-1}$  et directe  $H$ , s'appuient comme on l'a dit, sur la définition des modèles des caméras ayant un point de fuite  $P$  commun (en anglais : VIEW POINT OF THE PIN-HOLE MODEL).

Pour réaliser la transformation géométrique complexe  $H$ , on est amené à considérer en premier une opération appelée "projection-perspective"  $J$  qui permet de passer du point  $M$  d'une scène, repéré dans le repère terrestre ( $P_x, P_y, P_z$ ), au point  $m$  positionné dans le plan image d'une caméra réelle  $C_i$ , le point  $m$  étant alors repéré dans le système orthonormé individuel lié à cette caméra réelle, c'est-à-dire ( $O_{ixj}, O_{iyj}, O_{izj}$ ).

L'opération "projection-perspective"  $J$  fait passer, par une méthode mathématique, le repérage d'un point d'un système à 3 dimensions, dans un système à 2 dimensions. Cette opération peut être linéarisée par l'utilisation des coordonnées homogènes qui permet d'exprimer la transformation par une matrice  $4 \times 4$ .

La transformation géométrique  $H$  dite "transformation-perspective" est le produit de plusieurs matrices en coordonnées homogènes :

$T$  qui est une matrice représentant une translation,

$S$  qui est une matrice représentant un changement d'échelle,

$J$  qui est la matrice "projection-perspective" dont on a parlé plus haut,

$L$  qui est un changement d'origine,

$R$  qui est une rotation selon les angles  $\phi_j, \theta_j, \psi_j$ ,

ou  $\phi_o, \theta_o, \psi_o$  selon le cas.

La transformation géométrique directe s'exprime alors par :  $H = T * S * J * L * R$ , et la transformation géométrique inverse par :

$$H^{-1} = T^{-1} * S^{-1} * J^{-1} * L^{-1} * R^{-1}$$

Cette transformation peut être exprimée de manière approchée par un ensemble de fonctions polynomiales en utilisant par exemple la méthode des moindres carrés.

La transformation approchée ainsi obtenue peut alors être réalisée au moyen d'un produit du commerce référencé chez la firme TRW par TMC2302. L'adresse de cette firme est : PO Box 2472, LA JOLLA, CA92038 (US).

#### VI.b/ Description d'une mise en oeuvre particulière du calculateur d'adresse 200

Cette mise en oeuvre est illustrée au moyen de blocs fonctionnels sur la FIG.6, dans laquelle sont représentés:

- le bloc 201 qui est le premier module de construction du modèle de la caméra virtuelle MCo,
- le bloc 210 qui est le premier module de stockage des paramètres  $\phi_o, \theta_o, \psi_o, z_o$  de la caméra virtuelle imposés par l'utilisateur à l'interface 2 du bloc de contrôle 1,
- le bloc 113 qui est le générateur des adresses  $A_o$  de tous les pixels  $m'$  de l'image cible  $I_o$ ,
- le bloc 220 qui est un premier module de calcul de transformation géométrique qui reçoit les informations suivantes :
  - \* l'adresse  $A_o$  de  $m'$ , provenant du générateur d'adresse 113,
  - \* le modèle MCo de la caméra virtuelle provenant du module de construction 201,
  - \* les paramètres  $\phi_o, \theta_o, \psi_o, z_o$  de la caméra virtuelle provenant du premier module de stockage 210

Le bloc de calcul 220 détermine, dans l'espace à 3 dimensions, la position du rayon PM passant par le point de fuite  $P$  et par le pixel  $m'$ , à partir de la localisation de  $m'$  dans l'espace  $I_o$  à deux dimensions. Ce bloc de calcul 220 applique, à cet effet, la transformation perspective inverse :  $H_o^{-1}$  décrite précédemment,

- le bloc 204 qui est un module de stockage de la position du rayon lumineux PM dans l'espace à 3 dimensions, position qui est fournie par le bloc de calcul 220,
- le bloc de sélection 205 qui est un module d'aiguillage, et qui, à partir du rayon PM localisé, sélectionne la caméra réelle Cj, et fournit l'image source correspondante Ij dans laquelle trouver le point m recherché,
- le bloc 206 qui est le second module de construction des modèles MC1 à MCn des caméras réelles, avec des moyens incorporés de corrections des distorsions,
- les blocs 21.1 à 21.n qui sont les blocs de stockages, dits seconds moyens de stockage, des paramètres  $\phi_1$  à  $\phi_n$ ,  $\theta_1$  à  $\theta_n$ ,  $\psi_1$  à  $\psi_n$ ,  $z_1$  à  $z_n$  respectivement des caméras réelles,
- les blocs 22.1 à 22.n qui sont desdits seconds modules de calcul de transformation géométrique, et qui reçoivent respectivement les informations suivantes,
  - \* le signal de sélection provenant du module d'aiguillage 205, avec les informations relatives à l'image source Ij dans le cas décrit ici à titre d'exemple puisque la caméra Cj a été sélectionnée,
  - \* le modèle MCj de la caméra réelle Cj, provenant du second module de construction 206,
  - \* les paramètres  $\phi_j$ ,  $\theta_j$ ,  $\psi_j$ ,  $z_j$  de la caméra réelle Cj provenant du module de stockage 21.j (seconds moyens de stockage), dans le cas décrit ici. Dans le cas présent, le module de calcul 22.j (seconds moyens de calcul de transformation géométrique) calcule, dans l'espace à 2 dimensions constitué par le plan image source Ij, la localisation du point m, à partir de la localisation fournie par le module de stockage 204 du rayon lumineux PM, et ceci au moyen de la transformation perspective directe : Hj décrite précédemment (si le module d'aiguillage 205 avait sélectionné une autre image source, alors un autre module de calcul 22.1 à 22.n aurait appliqué la transformation perspective directe appropriée H1 à Hn pour localiser le point m dans cette autre image source),
- les blocs 23.1 à 23.n qui sont enfin des blocs de stockage des adresses Aq de chaque point m. Dans le cas présent l'adresse Aq du point m recherché sera disponible dans le module de stockage 23. j.

A la suite de ces opérations effectuées dans le calculateur d'adresse 200, en référence avec la Fig.3, ce calculateur d'adresse 200 fournit alors au bloc interpolateur 112 l'adresse Aq, et ce dernier

calcule la fonction de luminance F correspondante. L'image I0 peut alors être reconstituée dans le bloc de stockage 117 en faisant correspondre à chaque pixel m d'adresse Aq une valeur de la fonction F de luminance. L'image cible est maintenant reconstituée.

En référence avec la Fig.2, l'image I0 peut ensuite être affichée, ou stockée dans les moyens 10.

Si on se trouve dans le cas où plusieurs images sources sont possibles, par exemple parce qu'on a recouvrement des images sources dans une région frontière, le bloc de sélection 205 contient des moyens pour choisir une seule des deux images-sources.

Une méthode de sélection entre deux images consiste à décider que le raccord entre deux images dans la région frontière est constitué par une ligne droite qui passe sensiblement au milieu de la zone de recouvrement. Cette convention est applicable au cas des caméras juxtaposées horizontalement comme sur la Fig.1 avec des raccords verticaux.

L'homme du métier peut choisir toutes sortes de raccords. Par exemple des raccords disposés radialement sur des coniques.

En remarque, on peut noter que ce qui a été dit pour le codage des pixels en niveau de gris peut s'appliquer au codage d'une image couleur. Le codage s'effectue simplement sur un plus grand nombre de bits.

On notera que, pour qu'aucun raccord ne soit sensible dans l'image cible, il est nécessaire de prendre plusieurs mesures :

- la calibration précise des caméras fixes,
- la correction des résultats par des rebouclages, suivie d'une modification éventuelle des paramètres des modèles des caméras fixes jusqu'à ce qu'on obtienne une juxtaposition parfaite conduisant à une image cible sans défaut.

Cette opération, ou initialisation, doit être comprise comme une partie intégrante de l'opération de calibrage.

On rappelle en outre, que des corrections de distorsion sont à appliquer à chaque image source pour la construction des modèles MC1 à MCn. Il existe dans le commerce des dispositifs capables de réaliser ces opérations. La vérification de l'opération parfaite des correcteurs de distorsion utilisés est aussi partie intégrante du calibrage.

Par ailleurs, les différentes caméras réelles peuvent montrer des courbes de réponse en intensité de niveau de gris ou couleur qui diffèrent. Il est nécessaire lors du calibrage d'effectuer une correction de ces réponses individuelles pour uniformiser l'image virtuelle.

L'interface utilisateur 2 peut être commandée automatiquement, ou bien manuellement. Dans un exemple de mise en oeuvre cette interface peut être un levier orientable (en anglais JOY-STICK).

## Revendications

1. Procédé de traitement d'images, pour construire une image dite cible (Io), à partir d'images adjacentes, de cadrage fixe, dites images sources (I1,..., Ii, Ij,..., In), lesdites images cible et source ayant des points de fuite sensiblement confondus, **caractérisé en ce qu'il comprend** des étapes de :
  - numérisation des images,
  - détermination, pour un des pixels de l'image cible (Io), de l'adresse (Aq) d'un point dit correspondant, dans une image source (Ij) parmi toutes les images sources,
  - détermination de la valeur (F) de la luminance en ce point correspondant,
  - attribution au pixel de départ dans l'image cible (Io) de la valeur de luminance (F) de ce pixel correspondant,
  - et répétition de ces étapes pour chacun des pixels de l'image cible (Io).
2. Procédé selon la revendication 1, caractérisé en ce qu'il comprend une étape de calibrage incluant :
  - le stockage des paramètres des images sources, parmi lesquels les facteurs d'échelle ( $z1... zn$ ) et l'orientation des axes optiques concrétisée par les angles d'azimut ( $\theta1... \theta n$ ), de site ( $\phi1... \phi n$ ) et de rotation ( $\psi1... \psi n$ ) dans un repère orthonormé fixe, de référence (Px, Py, Pz),
  - la construction de modèles de caméras dites réelles pour fournir les images sources avec un point de fuite sensiblement commun confondu avec l'origine du repère de référence et avec aberrations corrigées, et en ce qu'il comprend une étape d'entrée de données incluant :
    - le stockage des paramètres de l'image cible parmi lesquels le facteur d'échelle ( $z0$ ) et l'orientation de son axe optique concrétisée par son angle d'azimut ( $\theta0$ ), de site ( $\phi0$ ) et rotation ( $\psi0$ ) dans le repère orthonormé fixe de référence (Px, Py, Pz),
    - la construction d'un modèle de caméra dite virtuelle pour fournir l'image cible avec un point de fuite sensiblement commun avec celui des images sources.
3. Procédé selon la revendication 2, caractérisé en ce qu'il comprend aussi des étapes de :
  - génération d'une adresse (Ao) pour chacun des pixels de l'image cible (Io), pixel par pixel, de manière à couvrir la totalité de l'image cible (Io),
  - détermination de l'image source particulière (Ij) et de l'adresse (Aq), dans cette image source, du point correspondant à chaque pixel de l'image cible (Io), à partir des paramètres des images source et cible.
4. Procédé selon la revendication 3, caractérisé en ce que la détermination de l'adresse (Aq) du point dit correspondant trouvé dans une image source (Ij) inclut :
  - une étape de détermination, à partir de l'adresse (Ao) d'un pixel de départ (m) dans l'image cible (Io), de l'orientation, dans le repère orthonormé fixe de référence (Px, Py, Pz) d'origine au point de fuite, du rayon lumineux passant par ce pixel donné (m) et par le point de fuite commun (P),
  - une étape de sélection de l'image source (Ij) traversée par ce rayon lumineux (Px),
  - et une étape de détermination de l'adresse (Aq), dans cette image source, du point d'intersection de ce rayon lumineux avec cette image source, ce point étant le point correspondant au pixel de départ.
5. Procédé selon la revendication 4, caractérisé en ce que :
  - l'étape de détermination de la position du rayon lumineux dans le repère de référence comprend une transformation géométrique de perspective dite "transformation de perspective inverse" ( $Ho^{-1}$ ) qui prend en compte le modèle de la caméra virtuelle,
  - l'étape de détermination de l'adresse (Aq) du point dit correspondant au pixel de départ, comprend des transformations géométriques de perspective ( $H1-Hn$ ) dites transformations de perspective directes, qui, selon l'image source particulière sélectionnée, prennent en compte respectivement le modèle de la caméra réelle correspondante.
6. Procédé selon l'une des revendications 1 ou 5, caractérisé en ce qu'il comprend une étape d' :
  - interpolation pour calculer une valeur d'une fonction de luminance (F) la plus

probable à l'adresse (Aq) trouvée l'image source (lj).

7. Dispositif de traitement d'images incluant :

- un système de n caméras réelles fixes (C1,... Ci, Cj, ... Cn), disposées de manière à ce que leurs champs de vue individuels fusionnent pour former un seul champ de vue de grand angle pour l'observation d'une scène panoramique, 5
- un système de construction d'image (100) simulant une caméra dite virtuelle (Co), mobile, balayant continuellement la scène panoramique pour en fournir une sous-image dite cible (lo), correspondant à une section arbitraire du champ de vue de grand angle, et construite à partir des images dites sources (l1,...li, lj... ln) adjacentes fournies par les n caméras réelles, cette caméra virtuelle (Co) ayant un point de fuite (P) confondu avec, ou proche de celui des caméras réelles, **caractérisé en ce que** ce système de traitement d'image est numérique et en ce que le système de construction (100) de l'image cible lo inclut : 10
- un calculateur d'adresse (200), pour faire correspondre à une adresse (Ao) d'un pixel dans l'image cible (lo), un point à une adresse (Aq) dans une des images sources (l1-ln), 15
- des moyens (112, 117) pour calculer la valeur (F) de luminance du point à l'adresse (Aq) trouvée dans l'image source (l1-ln), et pour affecter cette valeur (F) de luminance au pixel de départ à l'adresse (Ao) dans l'image cible (lo). 20

8. Dispositif selon la revendication 7, caractérisé en ce que le système (100) de reconstruction de l'image cible (lo) comprend : 25

- des premiers moyens de stockage (210) des paramètres relatifs à la caméra virtuelle pour fournir au calculateur d'adresse (200) le facteur d'échelle (zo) et l'orientation de l'axe optique de la caméra virtuelle (Co) dans un repère terrestre orthonormé (Px, Py, Pz) fixe et indépendant des caméras, c'est-à-dire ses angles d'azimut ( $\theta_0$ ), de site ( $\phi_0$ ) et de rotation ( $\psi_0$ ) ; 30
- des seconds moyens de stockage (21.1-21.n) des paramètres relatifs aux caméras réelles (C1-Cn) pour fournir au calculateur d'adresse (200) le facteur d'échelle (z1-zn) et l'orientation de l'axe optique de chacune des caméras réelles (C1,...Ci, Cj, Cn) c'est-à-dire leurs angles 35

d'azimut ( $\theta_1 - \theta_n$ ), de site ( $\phi_1 - \phi_n$ ) et de rotation ( $\psi_1 - \psi_n$ ) dans ledit repère terrestre fixe ;

- un générateur d'adresse (113) pour générer les adresses (Ao) des pixels de l'image cible (lo), pixel par pixel, de manière à couvrir la totalité de l'image cible (lo), le calculateur d'adresse (200) déterminant alors l'image source particulière (Cj) et le point à l'adresse (Aq) dans cette image source, qui correspond à chaque pixel de l'image cible (lo) à partir des paramètres des caméras virtuelle et réelles. 40

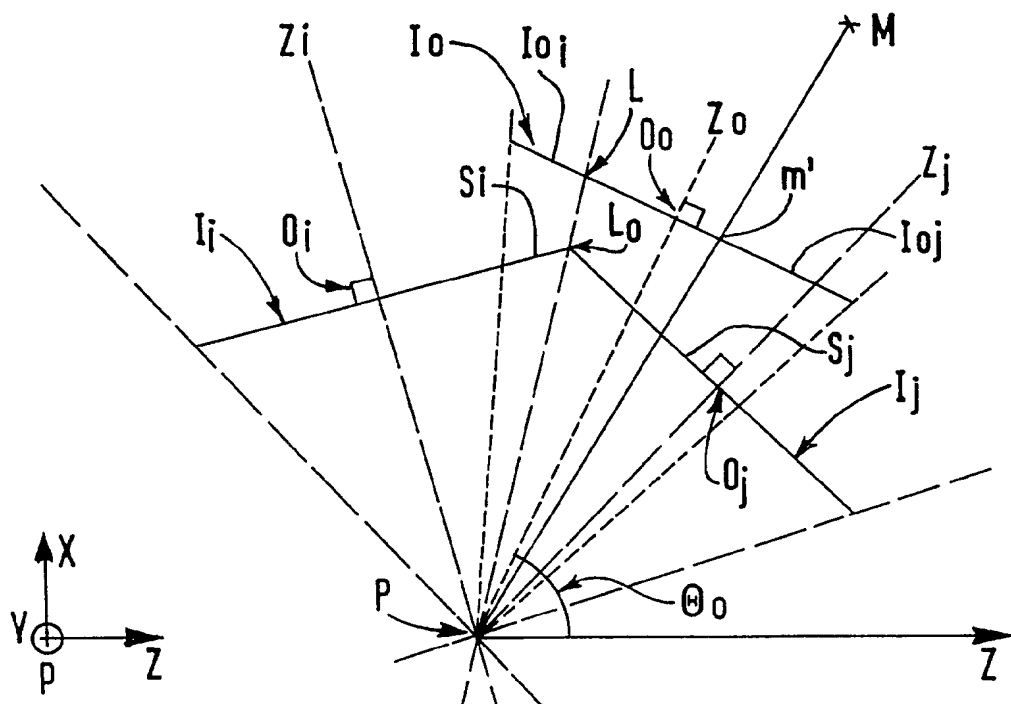
9. Dispositif selon la revendication 8, caractérisé en ce que le calculateur d'adresse (200) comporte : 45

- des premiers moyens de construction (201) d'un modèle (MCo) de caméra virtuelle avec projection à travers le point de fuite P, 50
- des seconds moyens de construction (206) de modèles (MC1-MCn) de caméras réelles, avec projection à travers le point de fuite P, et avec corrections des distorsions et défauts de perspective. 55

10. Dispositif selon la revendication 9, caractérisé en ce que le calculateur d'adresse (200) comporte : 60

- des premiers moyens de calcul de transformation géométrique (220) pour appliquer à chaque pixel (m) à une adresse (Ao) de l'image (lo) de la caméra virtuelle (Co), une transformation géométrique dite "transformation perspective" inverse ( $H_0^{-1}$ ), qui prend en compte le modèle de la caméra virtuelle MCo fourni par les premiers moyens de construction (201), et les paramètres angulaires d'azimut ( $\theta_0$ ), de site ( $\phi_0$ ), de rotation ( $\psi_0$ ) et le facteur d'échelle (zo) de cette caméra virtuelle, fournis par les premiers moyens de stockage (240), pour déterminer à partir de cette transformation perspective inverse ( $H_0^{-1}$ ), le positionnement dans ledit repère terrestre, du rayon lumineux passant par ce pixel (m) et le point de fuite (P), 65
- des moyens (204) pour stocker le positionnement du rayon lumineux obtenu par la transformation perspective inverse ( $H_0^{-1}$ ), 70
- des moyens de sélection (205) pour sélectionner l'image source particulière (l1-ln) traversée par ce rayon lumineux, 75

- des seconds moyens de calcul de transformation géométrique (22.1-22.n) pour appliquer à ce rayon lumineux dans ledit repère terrestre, une transformation géométrique dite "transformation perspective directe" (H1-Hn), qui prend en compte les modèles des caméras réelles fournis par les seconds moyens de construction (206), les paramètres angulaires d'azimut ( $\theta_1-\theta_n$ ) de site ( $\phi_1-\phi_n$ ), de rotation ( $\psi_1-\psi_n$ ), et le facteur d'échelle ( $z_1-z_n$ ) de la caméra (C1-Cn) réelle correspondante, fournis par les seconds moyens de stockage (21.1-21.n),
  - et des moyens de stockage (23.1-23.n) pour fournir, à partir de cette transformation perspective directe (H1-Hn), l'adresse (Aq) dans l'image source particulière (I1-In) qui correspond au rayon lumineux, et donc au pixel d'adresse (Ao) dans l'image cible (Io).
- 11.** Dispositif selon la revendication précédente 10, caractérisé en ce que les moyens de détermination de luminance comprennent :
- un interpolateur (112) pour calculer une valeur la plus probable d'une fonction de luminance (F) à l'adresse (Aq) trouvée par le calculateur d'adresse (200) dans l'image source fournie les moyens de sélection (205) ;
  - des troisièmes moyens de stockage (117) pour affecter la valeur (F) de luminance correspondant à l'adresse (Aq) trouvée dans l'image source, au pixel de départ dans l'image cible (Io) à l'adresse (Ao) fournie par le générateur d'adresse (113).
- 12.** Dispositif selon l'une des revendications précédentes, caractérisé en ce que le système (100) de reconstruction de l'image cible (Io) comprend aussi :
- une interface (2) pour permettre à un utilisateur de définir les paramètres de la caméra virtuelle (Co), ces paramètres incluant son facteur d'échelle ( $z_0$ ) et l'orientation de son axe optique ( $\theta_0, \phi_0, \psi_0$ ).
- 13.** Dispositif selon la revendication 12, caractérisé en ce que l'interface utilisateur (2) est commandée automatiquement ou manuellement.
- 14.** Dispositif selon l'une des revendications précédentes, caractérisé en ce que l'image cible (Io) est affichée en temps réel sur un écran (10).
- 15.** Dispositif selon l'une des revendications précédentes, caractérisé en ce que l'image cible (Io) est enregistrée.
- 16.** Dispositif de surveillance muni d'un dispositif de traitement d'image selon l'une des revendications 1 à 9.



**FIG. 1 B**

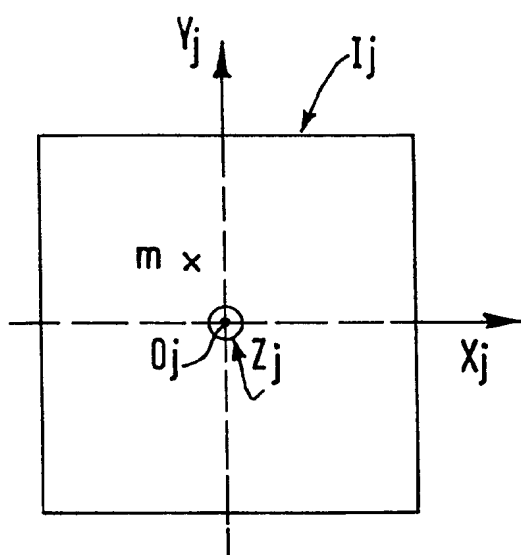
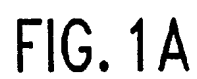


FIG. 1C

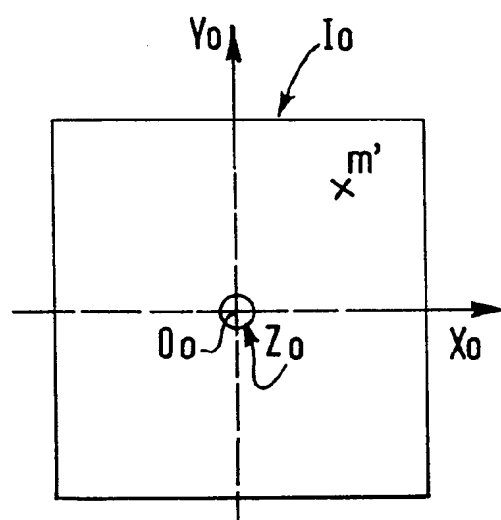


FIG. 1D



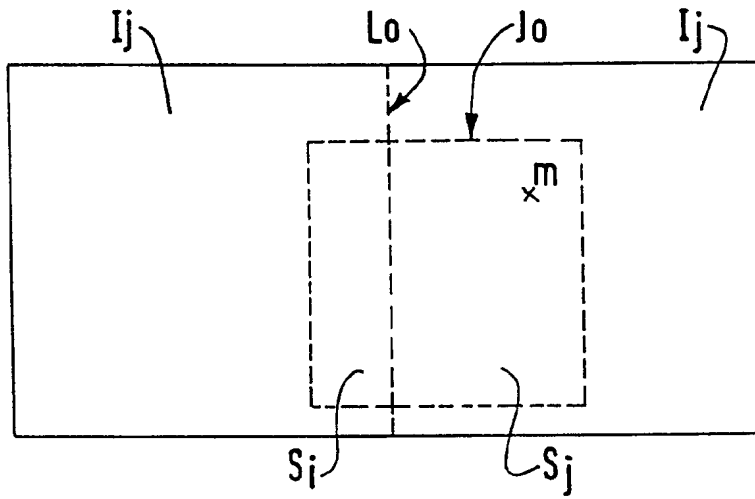


FIG. 1E

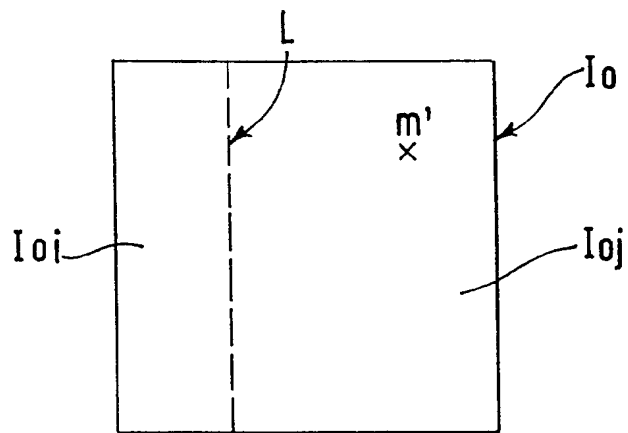
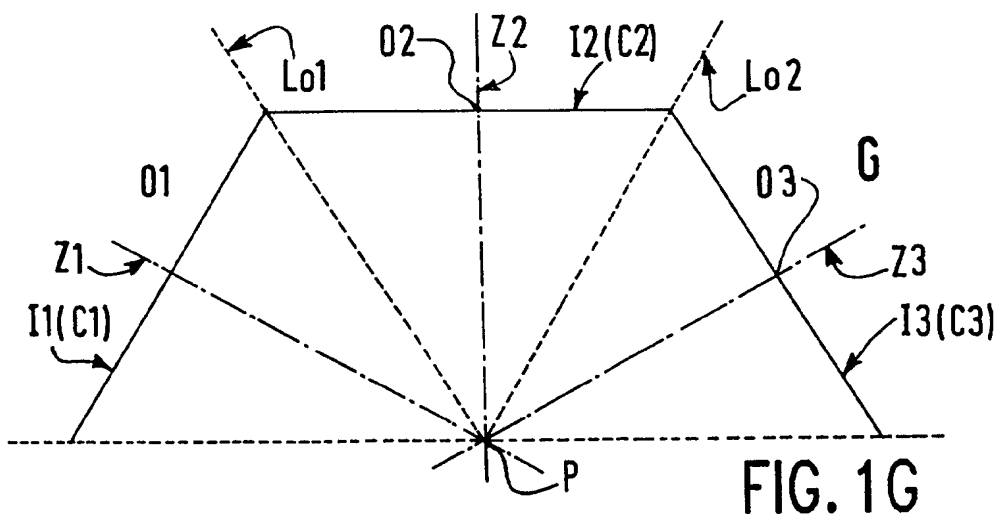
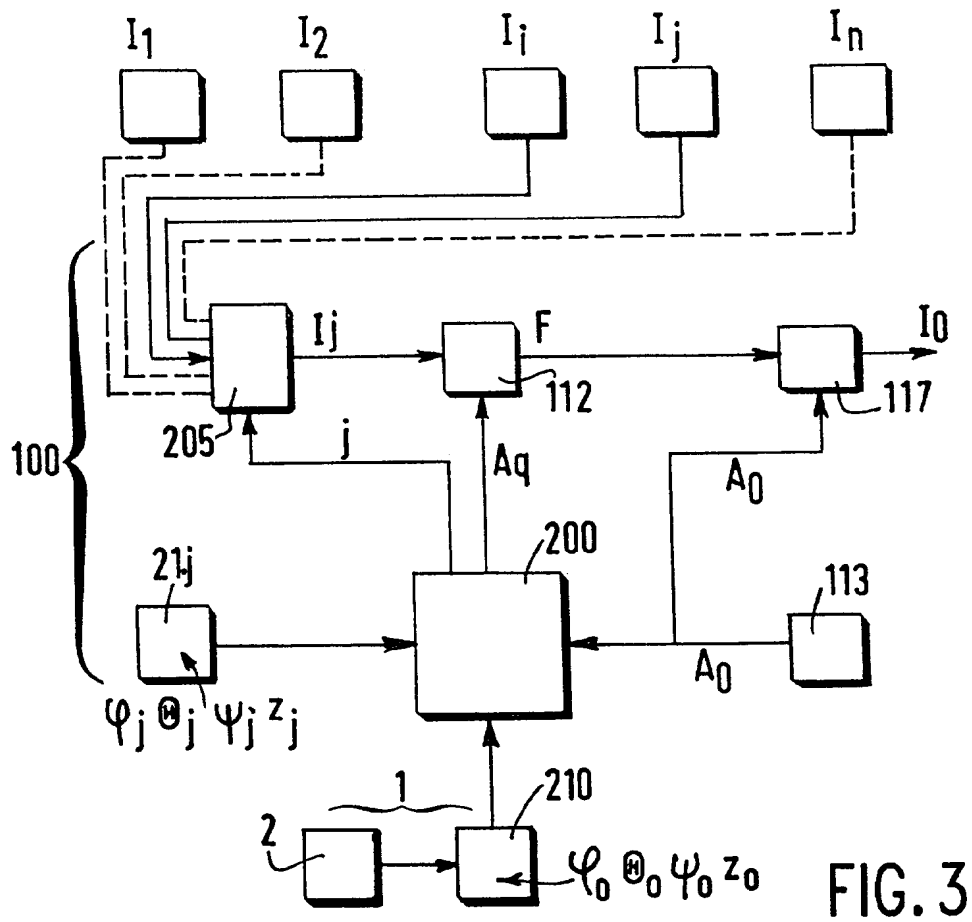
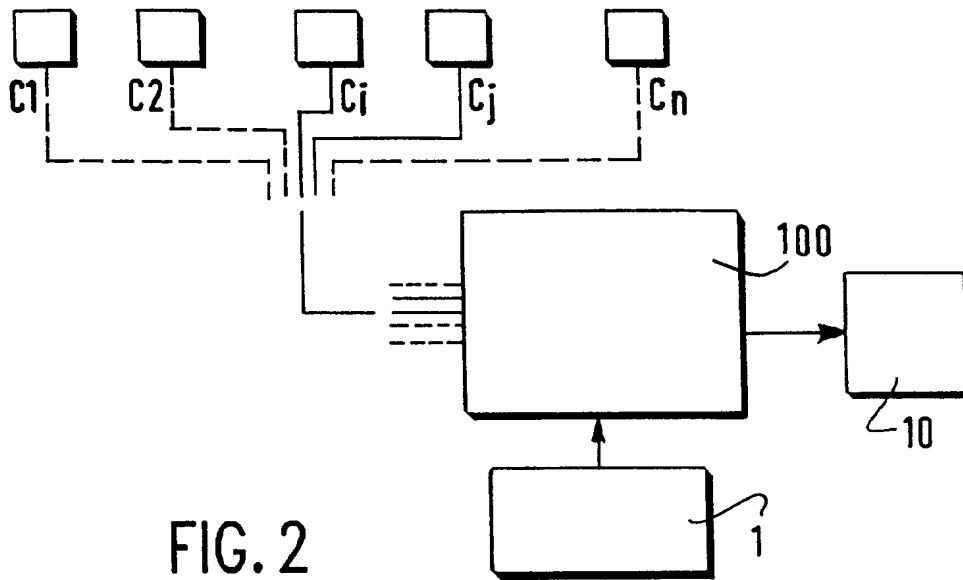


FIG. 1F





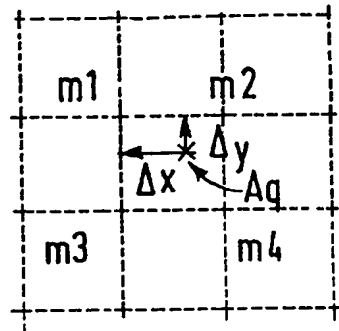


FIG. 4

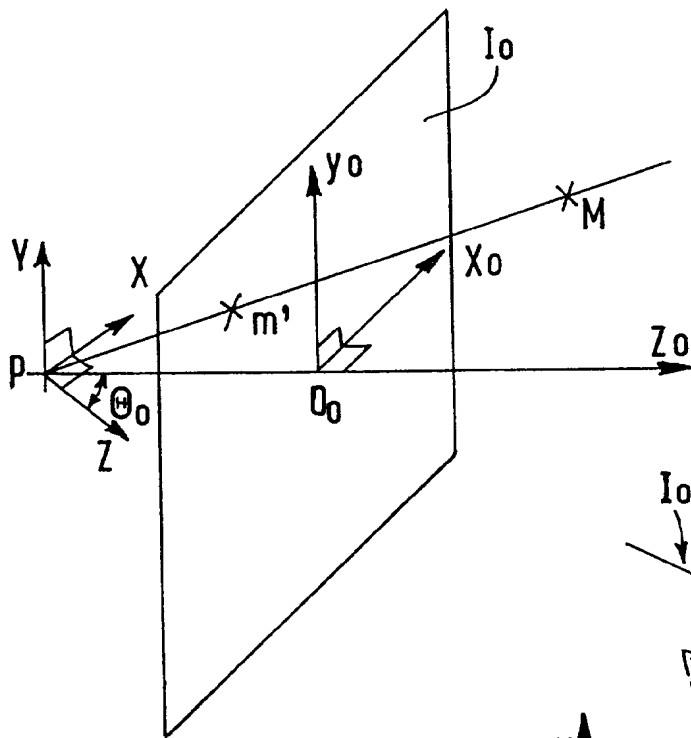


FIG. 5A

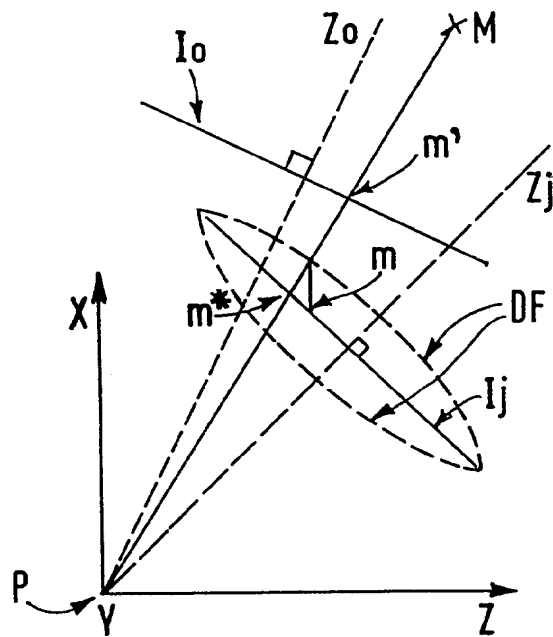


FIG. 5B

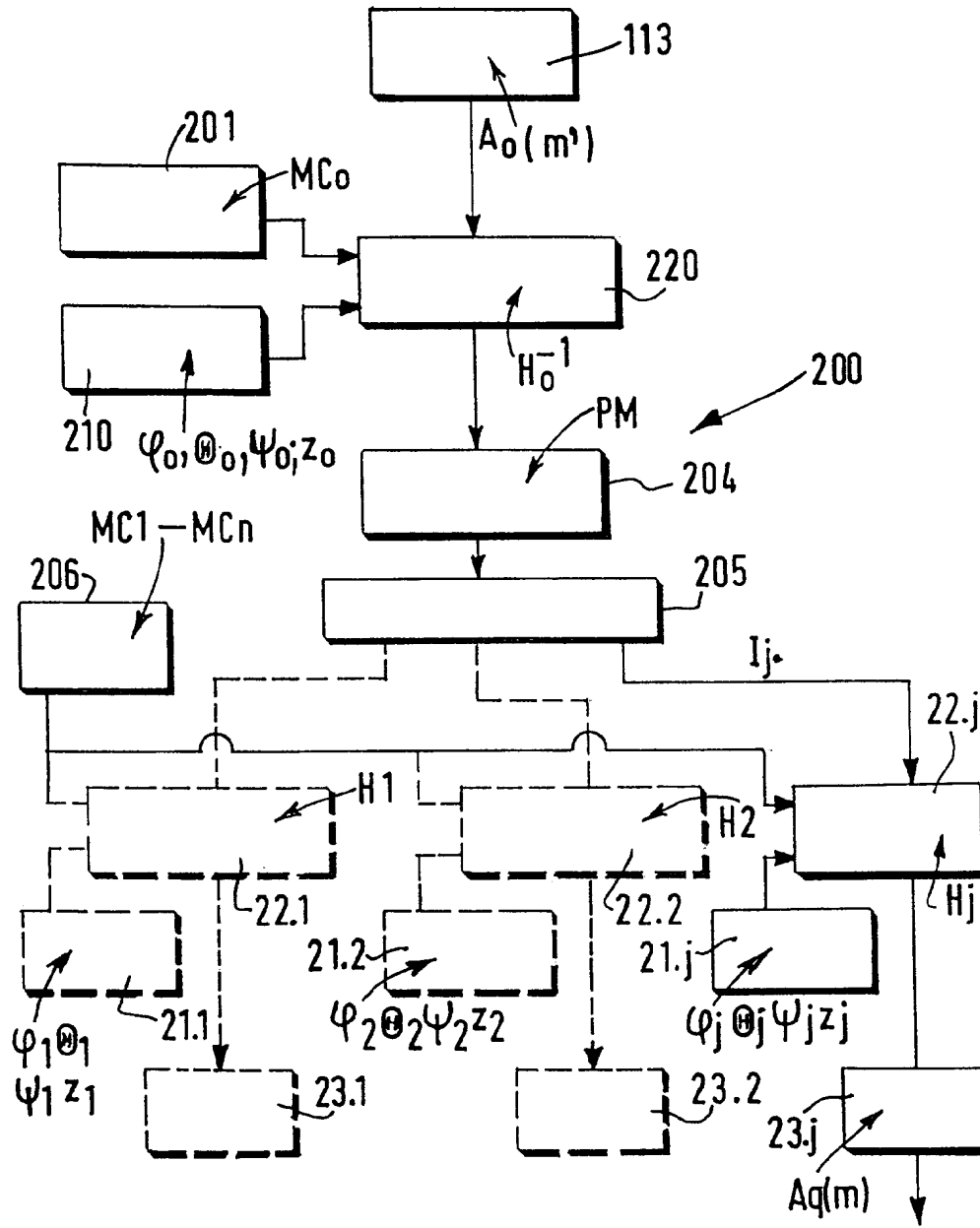


FIG. 6

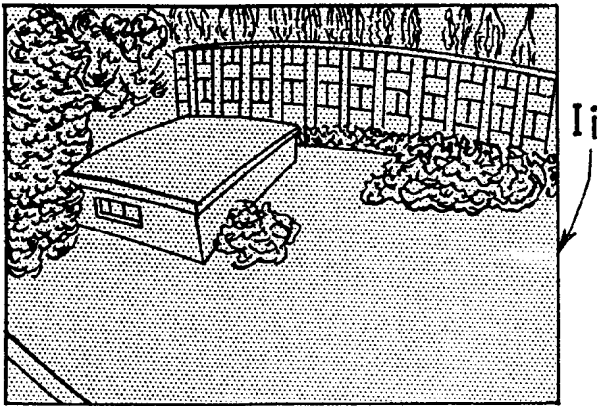


FIG. 7A

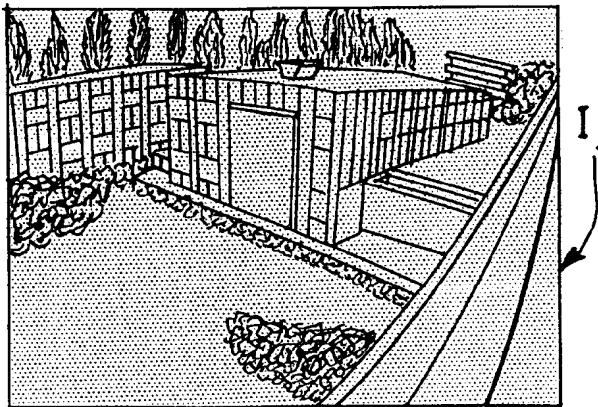


FIG. 7B

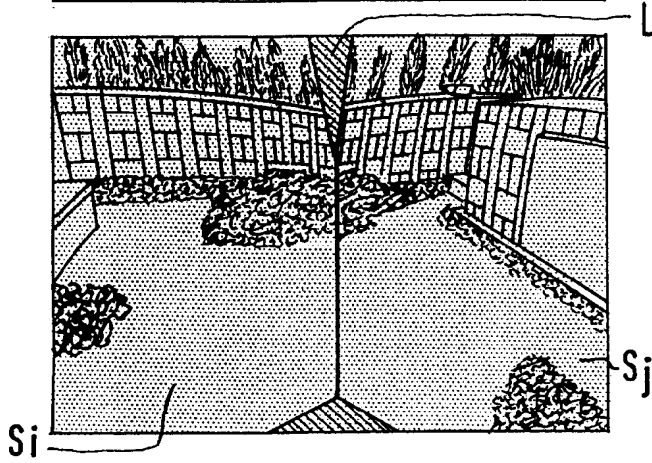


FIG. 7C

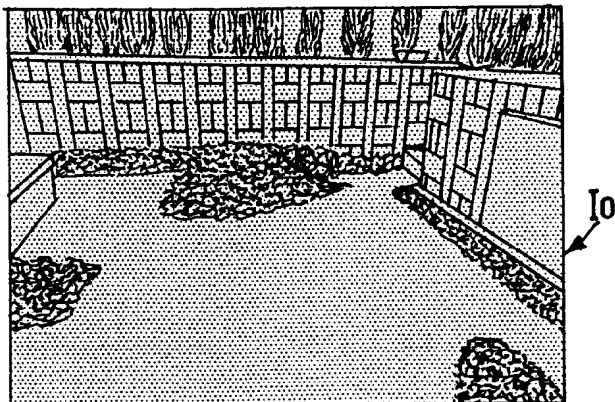


FIG. 7D



Office européen  
des brevets

# RAPPORT DE RECHERCHE EUROPEENNE

Numero de la demande  
EP 93 20 3620

DOCUMENTS CONSIDERES COMME PERTINENTS			
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (Int.Cl.5)
A	US-A-5 023 725 (MCCUTCHEN) 11 Juin 1991 * abrégé; figures 2,15-17,31 * * colonne 1, ligne 17 - colonne 2, ligne 14 * * colonne 16, ligne 50 - colonne 20, ligne 31 * ---	1-16	H04N7/18
A	US-A-4 677 576 (BERLIN, JR. ET AL.) 30 Juin 1987 * abrégé; figure 1 * * colonne 4, ligne 23 - colonne 5, ligne 43 * ---	1-16	
A	US-A-4 660 157 (BECKWITH ET AL.) 21 Avril 1987 * abrégé; figure 1 * * colonne 1, ligne 23 - colonne 2, ligne 45 * * colonne 5, ligne 26 - colonne 13, ligne 51 * ---	1-16	
A	WO-A-86 01592 (HUGHES AIRCRAFT COMPANY) 13 Mars 1986 ---	1-16	DOMAINES TECHNIQUES RECHERCHES (Int.Cl.5) H04N
A	EP-A-0 238 340 (RANK PULLIN CONTROLS LIMITED) 23 Septembre 1987 * abrégé; figures 1-6 * * page 1, ligne 14 - page 2, ligne 2 * * page 3, ligne 9 - page 5, ligne 6 * ---	1-16	
D,A	US-A-4 740 839 (PHILLIPS) 26 Avril 1988 * abrégé; figures 2,3 * * colonne 1, ligne 12 - colonne 2, ligne 3 * -----	1,7	
Le présent rapport a été établi pour toutes les revendications			
Lieu de la recherche LA HAYE		Date d'achèvement de la recherche 17 Février 1994	Examineur Machwirth, C
<div>CATEGORIE DES DOCUMENTS CITES</div> <div><div>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire</div><div>T : théorie ou principe à la base de l'invention E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date D : cité dans la demande L : cité pour d'autres raisons ----- &amp; : membre de la même famille, document correspondant</div></div>			